UNIVERSITÉ DE MONTRÉAL

CENTRE DE RECHERCHE SUR LES TRANSPORTS

CASE POSTALE 6128, SUCCURSALE A, MONTRÉAL, QUÉBEC, H3C 3J7

EXTRACTION OF LOGICAL RELATIONS AND DUALITY GAP

by

Brigitte Jaumard* and François Soumis**

* École Polytechnique de Montréal.

** École Polytechnique de Montréal et Centre de recherche sur les transports, Université de Montréal.

Cette étude a été publiée grâce à une subvention du fonds F.C.A.R. pour l'aide et le soutien à la recherche.

Université de Montréal
Centre de recherche sur les transports - Publication #458-A
January 1988

Dépôt légal - 1er trimestre 1988

ABSTRACT

We study the determination of the permanent of a system of 0-1 linear constraints, i.e. the set of variables with a fixed value in every feasible solution. First we recall how one can extract logical relations by using surrogate constraints in a system of 0-1 linear constraints. In the case of variables with a fixed value (unary relations), we show that the efficiency of the method depends on the value of the duality gap of an optimization problem deduced from the system of 0-1 linear constraints. The results obtained are illustrated on a voting confidentiality problem which can be formulated either as an extension of a generalized assignment problem or as a partitioning problem. We show that, depending on the formulation, the number of variables of the permanent one can identify changes and that the differences can be explained in terms of duality gap.

Key words: 0-1 linear optimization; logical conclusions; duality gap.

- - - - - - - - - -

RÉSUMÉ

Ce papier s'intéresse à la détermination des variables à valeur fixe dans les systèmes de contraintes linéaires en variables 0-1. On rappelle comment extraire des relations logiques dans un système de contraintes linéaires en variables 0-1. Dans le cas particulier de l'identification de variables à valeur fixe, on montre que l'efficacité de la méthode dépend du saut de dualité d'un problème d'optimisation défini à partir du système de contraintes. On illustre les propriétés obtenues sur un problème de décodage de vote qui se formule soit comme une extension du problème d'affectation généralisée soit comme un problème de partitionnement. On montre que, suivant la formulation retenue, le nombre de variables à valeur fixe identifiées varie et que les différences obtenues s'expliquent en termes de saut de dualité.

Mots-clés : optimisation linéaire 0-1; conclusions logiques; saut de dualité.

1. INTRODUCTION

Consider the following system of linear inequalities in 0-1 variables :

$$\sum_{j=1}^{n} a_{ij}x_j \leq b_i \qquad i=1,2,\ldots,m \qquad (1)$$

where : - $A = (a_{ij})$ $1\leq i\leq m$, $1\leq j\leq n$ is an m x n matrix with real coefficients,

- $b\in R^m$

- $x\in\{0,1\}^n$

Assuming that system (1) has a solution, we establish a measure of the difficulty of determining whether a variable has a fixed value for all the solutions of the system (1). The set of all variables taking fixed values will be called the "permanent of (1)", thus generalizing the definition of permanent given for the covering problem in graphs (see Hammer, Hansen and Simeone, 1982). To identify all the elements of the permanent is an NP-complete problem (Valiant, 1976).

To identify the elements of the permanent of (1), we shall use a method based on the extraction of logical relations using surrogate constraints (Jaumard and Minoux, 1984). In the next paragraph, we recall the main elements of this technique. We then show how its efficiency in determining whether a variable x_k takes a fixed value depends on the duality gap of two problems (P_k) and (Q_k) corresponding respectively to: Maximize x_k and, Maximize $y_k = \bar{x}_k$ on a set of constraints to be later defined.

We will then present an illustration of this result on an application concerning problems of voting confidentiality where the maximum amount of information is to be given while insuring that no voting element

can be determined, that is the permanent of the system of constraints is empty. In this example, we will show in particular that, by a reformulation of the problem we can :

i) reduce the duality gaps of the problems (P_k) and (Q_k),

ii) easily calculate the permanent of the system of inequalities considered using techniques for the extraction of logical relations without recourse to much more laborious implicit enumeration techniques.

2. THE EXTRACTION OF LOGICAL RELATIONS IN A 0-1 LINEAR SYSTEM OF INEQUALITIES

When solving mathematical programming problems with 0-1 variables, it is often seen to be useful, before applying a general enumeration algorithm, to try to extract certain information in the form of logical relations between the variables :

- unary relations (a variable is fixed at 1 or 0),
- binary relations ($xy=0$, $\bar{x}y=0,\ldots,$ $x\bar{y}=0$ ou $\bar{x}\bar{y}=0$ with $\bar{x}=1-x$ and $\bar{y}=1-y$)
- ternary or higher order relations ($xyz=0$, $\bar{x}y\bar{z}\bar{t}=0,\ldots,$ etc.).

These relations can be combined either to show that a problem has no solution, or to reduce the problem (to eliminate or fix a certain number of variables) and thus to facilitate the solution process (eg. Hammer and Nguyen, 1972; Breu and Burdet, 1974; Hansen, 1974; Johnson, Kostreva and Suhl, 1985; Jaumard, 1986).

Simple and efficient techniques exist to extract logical conclusions from each constraint of the problem taken on its own. However, for some problems, the consideration of constraints one by one severely limits the amount of useful information which can be extracted from the set of constraints. Generalizing the "constraint pairing" technique of Hammer, Padberg and Peled (1975), Jaumard and Minoux (1984), (see also Charnes et

al. [1975]), proposed a new method for extracting logical relations based the construction of a surrogate constraint linking together all the constraints of the problem.

Note also that similar work has been carried out by Chvatal and Hammer (1977) to determine the existence of a unique constraint equivalent (in terms of its solution set) to a system of inequalities in 0-1 variables, thus generalizing the results of Bradley (1971) and Rosenberg (1974) obtained for a system of linear equalities in 0-1 variables. This work appears to show that even when such a constraint exists, it is difficult to derive it. In contrast, the surrogate constraint suggested here is easy to determine but it is not equivalent to the system of inequalities (1).

In the following discussion, we recall the principle behind the extraction of logical relations on a surrogate constraint for a system of linear inequalities in 0-1 variables such as system (1).

We first transform the system into a system of equalities by associating a slack variable x_{n+i} with each inequality i,(i=1,2,...,m). An upper bound u_i is calculated for each slack variable. One possible value is $u_i = b_i - \sum_{j=1}^{n} \min(0, a_{ij})$. Other better bounds may be calculated by solving knapsack problems for example, (see Hammer, Padberg and Peled (1975)). The following system is then obtained :

$$\begin{cases} \sum_{j=1}^{n} a_{ij}x_j + x_{n+i} = b_i & i=1,2,\ldots,m \\ 0 \leq x_{n+i} \leq u_i & i=1,2,\ldots,m \\ x_j \in \{0,1\} & j=1,2,\ldots,n \end{cases} \qquad (2)$$

where the coefficients a_{ij} (i=1,2,...,m, j=1,2,...,n) and b_i (i=1,2,...,m) are real numbers (not necessarily non-negative).

Denote by t the vector in $\mathbb{R}^{n+m}$ such that :

$$\begin{cases} 1 \leq j \leq n & t_j = 1 \\ n+1 \leq j \leq n+m & t_j = u_{j-n} \end{cases}$$

and redefine the matrix A as an m x (n+m) matrix such that for $1 \leq i \leq m$, and $n+1 \leq j \leq n+m$

$$a_{ij} = \begin{cases} 1 & \text{if } i = j-n \\ 0 & \text{if } i \neq j-n \end{cases}$$

System (2) can then be rewritten as follows :

$$\begin{cases} \sum_{j=1}^{n+m} a_{ij} x_j = b_i & i=1,2,\ldots,m \\ x_j \in \{0,1\} & j=1,2,\ldots,n \\ 0 \leq x_j \leq t_j & j=n+1,\ldots,n+m \end{cases} \qquad (3)$$

By associating weights $\lambda_1, \lambda_2, \ldots, \lambda_m$ (without restriction sign) to the different constraints of (3) we obtain the composite surrogate constraint :

$$\sum_{j=1}^{n+m} \left(\sum_{i=1}^{m} \lambda_i a_{ij} \right) x_j = \sum_{i=1}^{m} \lambda_i b_i \qquad (4)$$

Using the following notation :

$$x_j^{\alpha_j} = \begin{cases} x_j & \text{if } \alpha_j = 1 \\ t_j - x_j & \text{if } \alpha_j = 0 \end{cases}$$

and expressing the terms with negative coefficients in (4) as a function of $x_j^0 = t_j - x_j$, we obtain the following equation with positive coefficients :

$$\sum_{j=1}^{n+m} \left| \sum_{i=1}^{m} \lambda_i a_{ij} \right| x_j^{\alpha_j} = \sum_{i=1}^{m} \lambda_i b_i - \sum_{j=1}^{n+m} t_j \min \left\{ 0, \sum_{i=1}^{m} \lambda_i a_{ij} \right\} \qquad (5)$$

Define the functions A_j, (j=1,2,...,n+m) by

$$A_j(\lambda) = \sum_{i=1}^{m} \lambda_i a_{ij}$$

and the function G by

$$G(\lambda) = \sum_{i=1}^{m} \lambda_i \, b_i - \sum_{j=1}^{n+m} t_j \min \{0, A_j(\lambda)\}$$

and, for all $S \subseteq \{1,2,\ldots,n\}$, the function F_S by :

$$F_S = G(\lambda) - \sum_{j \in S} \left| A_j(\lambda) \right| .$$

We then obtain :

Property 1 : If there exists a vector $\lambda = (\lambda_i) \in \mathbb{R}^m$ such that $G(\lambda) < 0$, then the system (1) has no solution.

Property 2 : If there exists a vector $\lambda = (\lambda_i)$ $i=1,2,\ldots,m \in \mathbb{R}^m$ such that $F_S(\lambda) < 0$, then for all solutions of (1): $\prod_{j \in S} x^{\alpha_j} = 0$.

It has been proved (Jaumard and Minoux, 1985) that to establish whether there exists such a vector λ, it is sufficient to evaluate the function F_S at a finite number of vectors corresponding to the extreme points and rays of a given system of linear equations.

In the general case, the functions F_S and G are continuous non-differentiable functions and the function G is convex. In the particular case where S reduces to a single element k, the corresponding function $F_k = F_{\{k\}}$ is equal to the minimum of two convex functions :

$$\text{for all } \lambda \in \mathbb{R}^m \qquad (F_k(\lambda) = \min (F_k^+(\lambda), F_k^-(\lambda))$$

where :

$$\begin{cases} F_k^+(\lambda) = G(\lambda) - A_k(\lambda) \\ F_k^-(\lambda) = G(\lambda) + A_k(\lambda) \end{cases}$$

with F_k^+ and F_k^- convex.

It is easy to show that for $|S| \neq 1$, the function F_S can be expressed as the minimum of $2^{|S|}$ convex functions. For $|S| = 2$, the derivation of logical conclusions (fixed variables or inconsistency or identification) from logical relations of the type $\prod_{j \in S} x^{\alpha_j} = 0$ can be done in

polynomial time (see e.g. Hansen, Jaumard and Minoux, 1985), while for $S > 3$ the problem is NP-complete (see Cook, 1971). However, given the increase in the number of functions studied (four for $|S| = 2$), we restrict the discussion to the case $|S| = 1$ from now on.

Non-differentiable optimization techniques, in particular the subgradient methods, will allow in practice to determine whether properties 1 and 2 are true.

Property 3 : If system (1) has a solution X and if there exists a vector $\lambda \in \mathbb{R}^m$ such that

$$F_k(\lambda) < 0, \text{ then :}$$

$$\text{if} \quad A_k(\lambda) > 0, \text{ then } x_k = 0 \text{ and } F_k(\lambda) = F_k^+(\lambda)$$

$$\text{otherwise } x_k = 1 \text{ and } F_k(\lambda)^- = F_k(\lambda).$$

Property 4 : If the system (1) has a solution and if there exists a vector $\tilde{\lambda}$, $\tilde{\lambda} \in \mathbb{R}^m$ such that :

i) $F_k^+(\tilde{\lambda}) < 0 \quad$ then $\quad F_k^-(\tilde{\lambda}) > 0$

ii) $F_k^-(\tilde{\lambda}) < 0 \quad$ then $\quad F_k^+(\tilde{\lambda}) > 0$

Proof : Since it is assumed that (1) has a solution, then for all $\lambda \in \mathbb{R}^m$,

$$F_k^+(\lambda) + F_k^-(\lambda) = 2G(\lambda) > 0$$

Thus there exists a vector $\tilde{\lambda} \in \mathbb{R}^m$ such that :

i) $F_k^+(\tilde{\lambda}) < 0$, then $\quad F_k^-(\tilde{\lambda}) = 2G(\lambda) - F_k^+(\tilde{\lambda}) > 0$

ii) $F_k^-(\tilde{\lambda}) < 0$, then $\quad F_k^+(\tilde{\lambda}) = 2G(\lambda) - F_k^-(\tilde{\lambda}) > 0.$

Q.E.D.

3. LINKS BETWEEN FIXED VARIABLES AND THE DUALITY GAP

The method of extraction of logical relations presented in Section 2 does not always result in the identification of all variables taking fixed values in all the solutions of the system (1). In this section we

present a characterization of the variables taking fixed values in all solutions with respect to the surrogate constraint method.

Let x_k be a binary variable taking a fixed value in all solutions of the system (1). We will show that the existence of a vector λ, $\lambda \in \mathbb{R}^m$, such that $F_k(\lambda) < 0$ depends on the duality gap of the following linear programs:

$$(P_k)\begin{cases} \text{Maximize } x_k \\ \text{subject to} \begin{cases} \sum_{j=1}^{n+m} \alpha_{ij} x_j = b_i & i=1,2,\ldots,m \\ 0 \leq x_j \leq t_j & j=1,2,\ldots,n+m \quad j\neq k \\ 0 \leq x_k \end{cases} \end{cases}$$

in the case where $x_k = 0$, and

$$(Q_k)\begin{cases} \text{Maximize } \bar{x}_j \\ \text{subject to} \begin{cases} \sum_{\substack{j=1 \\ j\neq k}}^{n+m} \alpha_{ij} x_j = \alpha_{ik}\bar{x}_k = b_i - a_{ik} & i=1,2,\ldots,m \\ 0 \leq x_j \leq t_j & j=1,2,\ldots,n+m \quad j\neq k \\ 0 \leq \bar{x}_k \end{cases} \end{cases}$$

in the case where $x_k = 1$.

Theorem 1 : Let x_k be a binary variable taking the value 0 in all solutions of the system (1).

There exists a vector λ, $\lambda \in \mathbb{R}^m$, such that $F_k(\lambda) < 0$ if and only if the duality gap of (P_k) is strictly less than 1.

Proof :

Let (DP_k) be the dual of (P_k).

$$(DP_k)\begin{cases} \text{Minimize } \sum_{i=1}^{m} \lambda_i b_i + \sum_{\substack{j=1 \\ j\neq k}}^{m} \mu_j t_j \\ \text{subject to} \begin{cases} \sum_{i=1}^{m} \lambda_i a_{ij} + \mu_j \geq 0 & \text{for all } j\neq k \\ \sum_{i=1}^{m} \lambda_i a_{ik} \geq 1 \\ \mu_j \geq 0 & \text{for all } j\neq k \end{cases} \end{cases}$$

We deduce that :

$$\mu_j = \max \{0, -\sum_{i=1}^{m} \lambda_i a_{ij}\} \qquad \text{for all } j\neq k$$

(DP_k) can thus be rewritten as a linear program with only one constraint:

$$\begin{cases} \text{Minimize } \sum_{i=1}^{m} \lambda_i b_i - \sum_{\substack{j=1 \\ j\neq k}}^{n+m} t_j \min (0, \sum_{i=1}^{m} \lambda_i a_{ij}) \\ \text{Subject to : } \sum_{i=1}^{m} \lambda_i a_{ik} \geq 1 \end{cases}$$

As $\sum_{i=1}^{m} \lambda_i a_{ik} \geq 1$ implies $\min (0, \sum_{i=1}^{m} \lambda_i a_{ik}) = 0$, the above program can be rewritten :

$$\begin{cases} \text{Minimize } G(\lambda) \\ \text{Subject to : } \sum_{i=1}^{m} \lambda_i a_{ik} \geq 1 \end{cases}$$

Let x^* be the optimal vector of the problem (P_k), and let λ^* be the optimal vector of the problem (DP_k).

<u>Necessary condition</u> : if the duality gap is strictly less than 1, then

$$G(\lambda^*) < 1 \quad \text{with} \quad \sum_{i=1}^{m} \lambda_i^* a_{ik} > 1$$

Hence $F_k^+(\lambda^*) = G(\lambda^*) - \sum_{i=1}^{m} \lambda_i^* a_{ik} < 0$

and $F_k(\lambda^*) = \min (F_k^+(\lambda^*), F_k^-(\lambda^*)) < 0.$

Sufficient condition :

If there exists a λ such that $F_k(\lambda) < 0$, as the variable x_k is fixed at 0,

$$F_k(\lambda) = F_k^+(\lambda) \quad \text{et} \quad \sum_{i=1}^{m} \lambda_i a_{ik} > 0$$

Thus $G(\lambda) - \sum_{i=1}^{m} \lambda_i a_{ik} < 0$

and $G(\lambda) < \sum_{i=1}^{m} \lambda_i a_{ik}.$

Define $\bar{\lambda}$ so that $\bar{\lambda}_j = \lambda_j / \sum_{i=1}^{m} \lambda_i a_{ik}$ (j=1,2,...,m). By linearity, we obtain :

$$G(\bar{\lambda}) < \sum_{i=1}^{m} \bar{\lambda}_i a_{ik}.$$

As $\sum_{i=1}^{m} \bar{\lambda}_i a_{ik} = 1$, $\bar{\lambda}$ is a feasible solution of (DP_k),

$$G(\lambda^*) \leq G(\bar{\lambda}) < \sum_{i=1}^{m} \bar{\lambda}_i a_{ik} = 1$$

an the duality gap is strictly less than 1. Q.E.D.

Theorem 2 : Let x_k be a binary variable taking the value 1 in all solutions of 1.

Then there exists a vector λ, $\lambda \in \mathbb{R}^m$ such that $F_k(\lambda) < 0$ if and only if the duality gap of Q_k is strictly less than 1.

Proof : Let (DQ_k) be the dual of (Q_k).

$$(DQ_k)\begin{cases} \text{Minimize } \sum_{i=1}^{m} \lambda_i\, b_i + \sum_{\substack{j=1 \\ j\neq k}}^{n+m} \nu_j\, t_j - \sum_{i=1}^{m} \lambda_i a_{ik} \\ \text{Subject to :} \begin{cases} \sum_{i=1}^{m} \lambda_i\, a_{ij} + \nu_j \geq 0 & \text{for all } j\neq k \\ -\sum_{i=1}^{m} \lambda_i\, a_{ik} \geq 1 & \text{for all } j\neq k \\ \nu_j \geq 0 \end{cases} \end{cases}$$

We therefore deduce that :

$$\nu_j = \max\left(0, -\sum_{i=1}^{m} \lambda_j\, a_{ij}\right) \quad \text{for all } j\neq k$$

(DQ_k) can be rewritten as linear program with a single constraint:

$$\begin{cases} \text{Minimize } \sum_{i=1}^{m} \lambda_i\, b_i - \sum_{\substack{j=1 \\ j\neq k}}^{n+m} t_j \min\left(0, \sum_{i=1}^{m} \lambda_i\, a_{ij}\right) - \sum_{i=1}^{m} \lambda_i\, a_{ik} \\ \text{Subject to : } \sum_{i=1}^{m} \lambda_i\, a_{ik} \leq -1. \end{cases}$$

As $\sum_i \lambda_i\, a_{ik} \leq -1$ implies that $\min\left(0, \sum_{i=1}^{m} \lambda_i\, a_{ik}\right) = \sum_{i=1}^{m} \lambda_i\, a_{ik}$, the above program can be rewritten :

$$\begin{cases} \text{Minimize } G(\lambda) \\ \text{Subject to : } \sum_{i=1}^{m} \lambda_i\, a_{ik} \leq -1. \end{cases}$$

Let x^* be the optimal vector of the problem (Q_k), and λ^* the optimal vector of the problem (DQ_k).

<u>Necessary condition</u> : if the duality gap is strictly less than 1,

$$G(\lambda^*) < 1 \quad \text{avec} \quad \sum_{i=1}^{m} \lambda_i^*\, a_{ik} \leq -1$$

Hence $$F_k^-(\lambda^*) = G(\lambda^*) + \sum_{i=1}^{m} \lambda_i^* a_{ik} < 0$$

and $$F_k(\lambda^*) = \min (F_k^+(\lambda^*), F_k^-(\lambda^*)) < 0$$

Sufficient condition :

If there exists a λ such that $F_k(\lambda) < 0$, as the variable x_k is fixed at 1,

$$F_k(\lambda) = F_k^-(\lambda) \quad \text{and} \quad \sum_{i=1}^{m} \lambda_i a_{ik} < 0$$

Thus $$G(\lambda) + \sum_{i=1}^{m} \lambda_i a_{ik} < 0$$

and $$G(\lambda) < - \sum_{i=1}^{m} \lambda_i a_{ik}.$$

Define $\bar{\lambda}$ so that $\bar{\lambda}_j = -\lambda_j / \sum_{i=1}^{m} \lambda_i a_{ik}$ $(j=1,2,\ldots,m)$. By linearity, we obtain

$$G(\bar{\lambda}) < - \sum_{i=1}^{m} \bar{\lambda}_i a_{ik}$$

As $\sum_{i=1}^{m} \bar{\lambda}_i a_{ik} = -1$, $\bar{\lambda}$ is a feasible solution of (DQ_k)

$$G(\lambda^*) \leq G(\bar{\lambda}) < - \sum_{i=1}^{m} \bar{\lambda}_i a_{ik} = 1$$

and the duality gap is strictly less than 1. Q.E.D.

In some cases the problems (P_k) and (Q_k) have duality gaps greater than or equal to 1, while the variable x_k takes a fixed value in all integer solutions of the system (1). Let us see an example.

Consider the following system of equalities :

$$\begin{cases} x_1 + x_2 = 1 \\ x_2 + x_3 = 1 \\ x_1 + x_3 + x_4 = 1 \\ x_4 + x_5 = 1 \end{cases} \qquad x_i \in \{0,1\} \ (i=1,2,\ldots,5) \qquad (6)$$

It is easy to verify that the system (6) has a unique solution (0,1,0,1,0).

We now illustrate theorems 1 and 2 with the problems (P_k) and (Q_k) (k=4 and 5). We first examine the variable x_5 and the problems (P_5) and (Q_5) :

$$(P_5)\quad \text{Maximize } x_5$$

$$\text{Subject to :}\quad \begin{aligned} x_1 + x_2 &= 1\\ x_2 + x_3 &= 1\\ x_1 + x_3 + x_4 &= 1\\ x_4 + x_5 &= 1\\ 0 \le x_i \le 1 &\qquad i=1,2,3,4\\ 0 \le x_5 & \end{aligned}$$

$$(Q_5)\quad \text{Maximize } \bar{x}_5$$

$$\text{Subject to :}\quad \begin{aligned} x_1 + x_2 &= 1\\ x_2 + x_3 &= 1\\ x_1 + x_3 + x_4 &= 1\\ x_4 + \bar{x}_5 &= 0\\ 0 \le x_i \le 1 &\qquad i=1,2,3,4\\ 0 \le \bar{x}_5 & \end{aligned}$$

The optimum value of (P_5) is equal to $z_5^* = 1$ with an optimal solution equal to $(\frac{1}{2},\frac{1}{2},\frac{1}{2},0,1)$, which does not allow us to conclude that there exists a unique value (equal to 0) for x_5 in every solution of (6). The same remark applies to the problem (Q_5) whose optimum value is equal to $\bar{z}_5 = 1$ and whose optimal solution is (0,1,0,1,1). Note on the other hand that if we examine the variable x_4, we obtain $\bar{z}_4 = 0$ and we conclude that $x_4 = 1$ in every integer solution of (6). If we then re-examine x_5, we may conclude that $x_5 = 0$, and similarly by examining the other variables, we obtain that (0,1,0,1,0) is the only integer solution of the system (6).

This example shows that when the duality gap of problems (P_k) and (Q_k) is initially greater than 1, it can sometimes be reduced to a value strictly less than 1 by fixing other variables.

4. APPLICATION TO A PROBLEM OF VOTING CONFIDENTIALITY

We now apply the theoretical results of the previous paragraphs to a class of problems arising in the analysis of voting confidentiality and related to the generalized assignment problem.

We consider a voting system defined by a group of N_a shareholders j (each with a weight p_j) who elect one individual among a group of N_c candidates. We will assume that a shareholder can only vote for one candidate, and that for any given candidate i, there is at least one shareholder j who votes for i. At the end of the vote, each candidate has collected a weight P_i, and the candidate with the greatest weight is elected. Suppose that we know : (i) the exact number of shareholders n_j who voted for a given candidate i and, (ii) some limits on the weights P_i collected by each candidate as a function of a precision parameter $\varepsilon (P_i^-(\varepsilon) \leq P_i \leq P_i^+(\varepsilon))$. We then ask the question whether the informations (i) and (ii) are sufficient to determine certain elements of the vote, and if so which ones. Define the binary variables x_{ij}, $(i=1,2,\ldots,N_c,\ j=1,2,\ldots,N_a)$ as follows:

$x_{ij} = 1$ is shareholder j voted for candidate i,

$x_{ij} = 0$ otherwise.

The vote remains secret provided that the permanent of the system (S_ε) below is empty, i.e. no variable has a fixed value in all solutions in (S_ε) :

$$(S_\varepsilon)\begin{cases} \sum_{i=1}^{N_c} x_{ij} = 1 & j=1,2,\ldots,N_a \\ P_i^-(\varepsilon) \leq \sum_{j=1}^{N_a} p_j x_{ij} \leq P_i^+(\varepsilon) & i=1,2,\ldots,N_c \\ \sum_{j=1}^{N_a} x_{ij} = n_i & i=1,2,\ldots,N_c \\ x_{ij} \in \{0,1\} & i=1,2,\ldots,N_c,\ j=1,2,\ldots,N_a \end{cases}$$

The system of equations (S_ε) obtained corresponds to a highly constrained problem (even more so when ε is small). In fact, the feasible region in real numbers is small, which produces large duality gaps. Frequently, in spite of the small number of possible solutions, it is difficult to extract information.

We illustrate the correlation between the duality gap and the difficulty of detecting fixed value variables on one of the instances of the problem which was provided. It had $N_c = 6$ and $N_a = 26$, i.e. a system (S_ε) of 156 variables and 44 constraints. Furthermore, we consider the case where ε=0.0005, $P_i^-(\varepsilon) = P_i - \varepsilon$ and $P_i^+(\varepsilon) = P_i + \varepsilon$ $(i=1,2,\ldots,N_c)$.

Using an explicit enumeration method, we show that this problem has 9 solutions and 98 fixed value variables (see Table 1 below).

shareholders

candidates i \ j	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25
1	-	0	0	0	-	0	0	1	0	0	0	-	0	0	0	0	0	0	0	0	0	0	0	0	-
2	-	0	-	0	0	0	-	0	-	0	0	-	-	0	-	-	-	0	-	0	0	0	-	0	0	0
3	-	-	-	0	0	-	-	0	0	0	-	0	0	0	-	0	0	-	0	0	-	-	0	0	0	0
4	0	0	0	1	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	1	0
5	-	-	-	0	-	-	0	0	-	0	0	-	-	0	-	0	-	-	0	-	-	0	-	1	0	-
6	-	0	-	0	0	-	-	0	-	1	-	-	-	0	0	-	-	-	-	-	-	-	-	0	0	-

Table 1 - The permanent of the system (S_ε), ε = 0.0005

Using a subgradient algorithm (Shor, 1968), we then calculated the minimum of the functions F_k, ($k=1,2,\ldots, N_a \times N_c = 156$) within an iterative procedure. As soon as a variable x_k is declared as belonging to the permanent (S_ε), (i.e.there exists a vector λ such that $F_k(\lambda) < 0$), the system (S_ε) is reduced. It may happen that the first time a variable x_k is examined, it is not detected as belonging to the permanent, but on a second examination after some variables have been fixed, a vector λ may be found such that $F_k(\lambda) < 0$.

After applying this procedure to the system (S_ε), ε=0.0005, the fixed variables given in Table 2 are obtained.

		shareholders																								
	i/j	1	2	3	4	5	6	7	8	9	10	11	12	14	14	15	16	17	18	19	20	21	22	23	24	2
candidates	1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	0	-	-	-	-	-	-	-	0	0	-
	2	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	0	0	-
	3	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	0	0	-
	4	-	-	-	-	-	-	0	-	0	-	-	0	0	-	-	0	0	-	-	0	-	-	0	0	1	-
	5	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	1	0	-
	6	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	0	0	-

Table 2 - The set of fixed variables obtained with the functions F_k

Note that only 20 variables out of 98 have been fixed by this method. For the 78 variables that were not identified as fixed, we checked whether the duality gap for the problems (P_k) and (Q_k) were greater than equal to 1. We found that when the variables were not fixed by the algorithm, this was due to the limits imposed by Theorems 1 and 2 and not to the premature stopping of λ prior to reaching optimality.

We then examined the effect of the parameter ε on the duality gap of problems (P_k) and (Q_k) and on the number of variables that could be fixed. When ε is reduced below 0,0005, the duality gaps are barely reduced

and the number of variables that can be fixed hardly changes. When ε is increased, and especially when $\varepsilon = 0.5$, the duality gaps increase and the number of variables that can be fixed is reduced. Table 3 shows the evolution of the duality gap of P_k for 4 typical variables.

ε	$P_{3,16}$	$P_{6,24}$	$P_{1,20}$	$P_{3,18}$
0,00005	0,961	0,863	1,000	1,000
0,00050	0,961	0,863	1,000	1,000
0,00500	0,962	0,864	1,000	1,000
0,05000	0,973	0,877	1,000	1,000
0,50000	1,000	1,000	1,000	1,000

Table 3 : Duality gap as a function of ε

5. A REFORMULATION OF THE PROBLEM

In this section we present a reformulation of the system (S_ε) as a partitioning problem which allows to determine more easily the permanent of the system (S_ε).

We recall below a classical method (Geoffrion, 1974; Fischer, 1981) for the transformation of a general assignment type problem into a partitioning problem.

Consider the following sub-systems $(S_i)_\varepsilon$, $(i=1,2,\ldots,N_c)$ of the systems (S_ε) :

$$(S_i)_\varepsilon \begin{cases} P_i^-(\varepsilon) \leq \sum_{j=1}^{N_a} p_j x_{ij} \leq P_i^+(\varepsilon) \\ \sum_{j=1}^{N_a} x_{ij} = n_i \\ x_{ij} \in \{0,1\} \qquad j=1,2,\ldots,N_a \end{cases}$$

We calculate the sets E_i of integer solutions for each system $(S_i)_\varepsilon$ of equations, and we define the matrix A^i of dimension $N_a \times |E_i|$ so that each column vector of A^i corresponds to a feasible solution of E_i.

Consider the following matrix A :

$$A = \begin{array}{|c|c|c|c|} \hline A^1 & A^2 & \cdots & A^{N_c} \\ \hline 11\ldots 1 & & & \\ & 11\ldots 1 & & \\ & & \cdots & \\ & & & 11\ldots 1 \\ \hline \end{array}$$

Denote $N = \sum_{i=1}^{N_c} |E_i|$.

The system of partitioning type equations

$$(R_\varepsilon) \quad \begin{cases} Ay = 1 \\ y \in \{0,1\}^N \end{cases}$$

where A is a matrix of dimension $(N_c + N_a) \times N$ is equivalent to the system of general assignment type equations (S_ε) in the sense that there is a one-to-one correspondence between the set of solutions of (R_ε) and the set of solutions of (S_ε). Indeed, a solution x of the problem (S_ε) can be defined on a solution y of R_ε :

$$x_{ij} = \sum_{\ell \in E_i} y_\ell \, a_{j\ell} \qquad i=1,2,\ldots,N_c,\ j=1,2,\ldots,N_a \tag{7}$$

Similarly, a solution y of problem (R_ε) can be obtained from a solution x of (S_ε). Let i be such that ℓ is a column of the sub-matrix A^i. Then we obtain :

$$y_\ell = \sum_{j=1}^{N_a} x_{ij} \, a_{j\ell} \tag{8}$$

To determine the solutions of the system (S_ε), the following problem may be considered :

$$(T_\varepsilon)\begin{cases} A_y = 1 \\ y \in \{0,1\}^N \\ x_{ij} = \sum_{\ell \in E_i} y_\ell \, a_{j\ell} \qquad i=1,2,\ldots,N_c, \quad j=1,2,\ldots,N_a \end{cases}$$

We experimented the method of extraction of unary relations from a surrogate constraint (i.e. the identification of fixed variables) on both problems (R_ε) and (T_ε) for different values of ε. At the same time, we used an enumeration method to determine the exact permanent of each problem and its corresponding system (S_ε). The results are summarized in Table 4 below. The functions $P_i^-(\varepsilon)$ and $P_i^+(\varepsilon)$ $(i=1,2,\ldots,N_c)$ were defined as follows :

$$P_i^-(\varepsilon) = \left\lceil \frac{P_i + \varepsilon}{2\varepsilon} \right\rceil x\ 2\varepsilon - \varepsilon$$

$$P_i^+(\varepsilon) = \left\lceil \frac{P_i + \varepsilon}{2\varepsilon} \right\rceil x\ 2\varepsilon + \varepsilon \ .$$

The symbol $\lceil x \rceil$ signifies the smallest integer greater than or equal to x. In the case where $\varepsilon = 0.5 \times 10^{-\alpha}$, $\left\lceil \frac{P_i + \varepsilon}{2\varepsilon} \right\rceil x\ 2\varepsilon$ is the value of the vote for candidate i rounded to α decimals.

ε	N	NS	N_0	N_1	t_R	t_T	0	1
0,00050	19	1	13	6	100	100	130	26
0,00175	62	1	56	6	100	100	130	26
0,00200	71	3	60	2	100	100	117	14
0,00250	78	2	67	1	100	100	114	10
0,00400	134	9	106	1	100	100	88	5
0,00500	168	17	143	1	100	100	93	5
0,00700	214	47	141	1	100	100	68	4
0,00800	241	27	189	1	100	100	68	4
0,01000	306	16	271	1	100	100	82	4

Table 4 : Experimental results for the determination of the permament on the problems (R_ε) et (T_ε)

ε = value of the precision parameter

N = the number of variables in the partitioning problem R_ε

NS = the number of solutions to the partitioning problem R_ε

N_0 = the number of fixed variables taking the value 0 in R_ε

N_1 = the number of fixed variables taking the value 1 in R_ε

t_R = percentage of variables of the permanent of R_ε obtained by the method of extraction of unary relations.

t_T = percentage of variables of the permanent of T_ε obtained by the method of extraction of unary relations.

Note that the method of extracting unary relations allows the determination of the exact permanent of the different problems (R_ε). Then, based on the problems (T_ε), we can determine the exact permanent of the problems (S_ε). This means that the duality gaps of the different problems (P_k) and (Q_k) corresponding to the variables of the permanent are strictly less than 1, if not on the initial problem, then at least after simplifying the problem by fixing part of the permanent.

The two last columns of the table indicate the number of variables fixed at 0 and at 1 in the permanent of S_ε. These numbers generally decrease as ε increases. Thus the vote can be completely decoded for ε=0.0005 and 0.00175, and secrecy increases with ε. However, more than half the variables are still fixed and identifiable for ε=0.01.

Note also that the computation times obtained on a CYBER are of the order of one minute for the set of problems treated (N > 100) and seem to be comparable for the two methods used. We have not given precise times in Table 4 because of the largely experimental character of the algorithms used : the objective of this study was not a comparative analysis of computation times for the different approaches, but a precise illustration of the relationship between the duality gap and the difficulty of solving a

highly constrained problem. In particular, the algorithm for the extraction of unary relations should be easily improved, by simultaneously exploiting the sparseness of the constraint matrix when certain variables are fixed, and defining an ordering criterion for the examination of variables.

6. CONCLUSION

The theorems of Section 3 establish necessary and sufficient conditions under which the method of extracting logical relations identifies whether or not a variable takes a fixed value in all solutions. The vote decoding example illustrates the fact that the conditions concerning the duality gap may be satisfied or not depending on the formulation of the problem.

BIBLIOGRAPHY

BRADLEY, G.H., "Transformation of Integer Programs to Knapsack Problems", Discrete Math. 1 (1971) 29-45.

BREU, R. and BURDET, C.A., "Branch and Bound Experiments in Zero-one Programming", Mathematical Programming Study 2 (1974) 1-50.

CHARNES, A., GRANOT, D. and GRANOT, F., "On Improving Bounds for Variables in Linear Integer Programs by Surrogate Constraints", INFOR 13, 3 (1975) 260-269.

CHVATAL, V. and HAMMER, P.L., "Aggregation of Inequalities in Integer Programming", Annals of Discrete Mathematics 1 (1977) 145-162.

COOK, S.A. "The Complexity of Theorem Proving Procedures", Proceedings of the Third Annual Symposium on Theory of Computing (1971) 151-158.

FISCHER, M.L., "The Lagrangian Relaxation Method for Solving Integer Programming Problems", Management Science 27 (1981) 1-18.

GEOFFRION, A.M., "Lagrangian Relaxation and its Uses in Linear Programming", Mathematical Programming Study 2 (1974) 82-114.

GOFFIN, J.L., "On Convergence Rates of Subgradient Optimization Methods", Mathematical Programming 13 (1977) 329-347.

HAMMER, P.L., PADBERG, M.W. and PELED, U.N., "Constraint Pairing in Integer Programming", Infor 13,1 (1975) 68-81.

HAMMER, P.L. and NGUYEN, S., "A Partial Order on the Solution Space of Bivalent Programs, APOSS", CRM 163, Université de Montréal (1972).

HAMMER, P.L., HANSEN, P. and SIMEONE, B., "Vertices Belonging to all or to no Maximum Stable Sets of a Graph", SIAM Journal of Algebraic and Discrete Methods 3 (1982) 511-522.

HANSEN, P., "Programmes mathématiques en variables 0-1", thèse d'aggrégation, Université Libre de Bruxelles (1974).

HANSEN, P., JAUMARD, B. and MINOUX, M., "A Linear Expected-time Algorithm for Deriving all the Logical Conclusions Implied by a Set of Boolean Inequalities", Mathematical Programming 34, 223-231 (1986).

JAUMARD, B., "Extraction et utilisation des relations booléennes pour la résolution des programmes linéaires en variables 0-1", Thèse de doctorat, École Nationale Supérieure des Télécommunications (1986).

JAUMARD, B. and MINOUX, M., "Extraction of Logical Relations on Variables in Discrete Optimization - Generalization of Constraint Pairing and New Algorithms", Rapport technique, CNET/PAA/TIM/MTI/1456 (1984).

JOHNSON, E.L., KOSTREVA, M.M. and SUHL, U.H., "Solving 0-1 Integer Programming Problems Arising from Large Scale Planning Models", Operations Research 33, 4 (1985) 803-819.

ROSENBERG, I.G., "Aggregation of Equations in Integer Programming", Discrete Math. 10 (1974) 325-341.

SHOR, N.Z., "On the Rate of Convergence of the Generalized Gradient Method", Kibernetika 4, 3 (1968) 98-99.

VALIANT, L.G., "The Complexity of Computing the Permanent", Theorical Computer Science 8 (1979) 189-221.

ACKNOWLEDGEMENTS

We would like to thank Michel Sauvé for his suggestions on some details in the algorithm, and his help for the implementation of it.